

基于元胞蚂蚁算法的故障诊断

朱刚¹, 周政新¹, 马良²

ZHU Gang¹, ZHOU Zheng-xin¹, MA Liang²

1. 上海第二工业大学 电子与电气工程学院, 上海 201209

2. 上海理工大学 管理学院, 上海 200093

1. College of Electronic and Electric Engineering, Shanghai Second Polytechnic University, Shanghai 201209, China

2. College of Management, University of Shanghai for Science and Technology, Shanghai 200093, China

E-mail: zhugang@ee.sspu.cn

ZHU Gang, ZHOU Zheng-xin, MA Liang. Fault diagnosis based on cellular ant algorithm. Computer Engineering and Applications, 2008, 44(33): 224-226.

Abstract: A new kind of fault diagnosis based on cellular ant algorithm is presented. The method realizes classification of fault by near-neighborhood criteria of pattern recognition and cellular ant algorithm. The simulation results show that the method is efficient for fault diagnosis.

Key words: fault diagnosis; cellular ant algorithm; near-neighborhood criteria

摘要: 提出一种基于元胞蚂蚁算法的故障诊断方式。该方法利用模式识别中的近邻准则, 使用元胞蚂蚁算法实现故障的分类, 达到故障诊断的目的。最后用一实例进行仿真, 结果表明该方法拥有较强的可行性和实用性。

关键词: 故障诊断; 元胞蚂蚁算法; 近邻准则

DOI: 10.3778/j.issn.1002-8331.2008.33.068 文章编号: 1002-8331(2008)33-0224-03 文献标识码: A 中图分类号: TP181; O23

1 引言

故障诊断是指系统在一定工作环境下, 查明导致系统某种功能失调的原因或性质, 判断劣化状态发生的部位或部件, 以及预测状态劣化的发展趋势等^[1]。由于诊断对象的复杂性, 人们往往提出了大量的原始特征和故障假设进行故障识别, 这样会导致系统效率低, 满足不了实时的要求; 同时随着现代工业及科学技术的发展, 各种生产设备日趋大型化、自动化和智能化, 传统的诊断技术已不能适应现有设备故障诊断的需要。为此引入人工智能方法已成为人们近期的共识^[2-4]。专家系统、模糊理论、神经网络^[5]、遗传算法^[6, 8]、智能主体^[9-12]及它们相互之间和它们与其它信息处理技术的融合是人工智能手段的主要表现方法, 在诊断领域中, 它们受到越来越广泛的重视。

在应用环境中, 故障征兆和原因的数据收集往往有一定的困难, 难以确定征兆与状态的对应关系。为此利用改进的蚂蚁算法——元胞蚂蚁算法^[13]进行故障特征提取, 实现故障分类, 为实际判断故障提供参考。

2 近邻函数准则

故障分类可转化为欧氏空间中的聚类问题, 常采用样本距离作为分类指标, 但对于某些不确定问题会出现误判, 为此常

引入近邻函数准则^[14]。

对于数据集中的任何两个样本 y_i, y_j , 若 y_j 是 y_i 的第 I 个近邻, 则称 y_j 对 y_i 的近邻系数为 I 。类似地, 若 y_i 是 y_j 的第 K 个近邻, 则称 y_i 对 y_j 的近邻系数为 K 。如果 y_i 和 y_j 互为最近邻, 这时的近邻函数值最小, 等于零。若用 α_{ij} 表示 y_i 和 y_j 之间的近邻函数值, 则有:

$$\alpha_{ij} = I + K - 2$$

若在聚类过程中, y_i 和 y_j 被分在同一类, 那么此时的 y_i 和 y_j 是相互“连接”的。对于每一个这样的“连接”存在着一个相应的“连接”损失。算法中, “连接”损失规定为两个样本间的近邻函数值 α_{ij} 。为了避免出现只包含一个样本的聚类。规定样本本身的“连接”损失, 即 α_{ii} 为 $2N$ 或更大, 这里的 N 是样本总数。由于 I, K 总是小于或等于 $N-1$, 因此 y_i 和 y_j 之间近邻函数的最大值为:

$$\max \alpha_{ij} = N - 1 + N - 1 - 2 = 2N - 4, i \neq j$$

显然, $\alpha_{ii} > \max \alpha_{ij}, i \neq j$ 。

这样规定“连接”损失的好处是使密度较为接近的点容易聚成一类。总类内损失规定为:

$$L_{IA} = \sum_{i=1}^N \sum_{j=1}^N \alpha_{ij}$$

基金项目: 国家自然科学基金(the National Natural Science Foundation of China under Grant No.70471065); 上海市教育委员会科研重点项目(No.DZ106001); 上海市重点学科建设项目(No.T0502); 上海市教育委员会重点学科建设项目(No.J51801)。

作者简介: 朱刚(1963-), 男, 副教授(博士), 从事系统工程、智能优化研究; 周政新(1957-), 男, 教授, 主要研究方向为智能故障诊断, 人工智能; 马良(1964-), 男, 教授, 博士, 博士生导师, 从事系统工程, 智能优化研究。

收稿日期: 2007-12-10

修回日期: 2008-03-03

同一类中的 y_i 与 y_j , 由于存在着“连接”关系, 所以 α_{ij} 不等于零。而不同类的 y_i 和 y_j , 由于不存在“连接”关系, 所以 α_{ij} 等于零。因此可以对所有样本进行求和。

为了计算类间损失, 首先计算类与类之间的最小近邻函数值 $\gamma_{j\theta}$, 即计算 i 类中的每一点与 j 类中所有各点的近邻函数值, 并取其中的最小值为 $\gamma_{j\theta}$ 。显然 i 类与其余的 $C-1$ 类的最小近邻函数值为:

$$\gamma_i = \min_{\substack{j \in \theta \\ j \neq i}} \gamma_{ij}, \theta = \{1, 2, \dots, C\}$$

设 $\alpha_{i \max}$ 是 i 类中两点间的最大“连接”损失, $\alpha_{k \max}$ 是 k 类中两点间的最大“连接”损失。定义其它类对 i 类的类间损失为:

$$\beta_i = \begin{cases} -[(\gamma_i - \alpha_{i \max}) + (\gamma_i - \alpha_{k \max})] & \gamma_i > \alpha_{i \max}, \gamma_i > \alpha_{k \max} \\ \gamma_i + \alpha_{i \max} & \gamma_i \leq \alpha_{i \max}, \gamma_i > \alpha_{k \max} \\ \gamma_i + \alpha_{k \max} & \gamma_i \leq \alpha_{k \max}, \gamma_i > \alpha_{i \max} \\ \gamma_i + \alpha_{i \max} + \alpha_{k \max} & \gamma_i \leq \alpha_{i \max}, \gamma_i \leq \alpha_{k \max} \end{cases}$$

则总类间损失为

$$L_{IR} = \sum_{i=1}^C \beta_i$$

总损失为 $J_{NV} = L_{IA} + L_{IR}$, 分类的结果应使准则函数 J_{NV} 取最小值。

3 元胞蚂蚁算法

元胞自动机是一个时间、空间、状态离散的并行数学模型。元胞蚂蚁算法是基于元胞自动机原理的一种改良的蚂蚁算法, 它在基本蚂蚁算法中引入元胞、元胞空间、邻居及元胞的演化规则。对于标准的 TSP 问题, 可描述如下:

若给定城市元素的集合 $C = \{c_1, c_2, \dots, c_n\}$, 则 C 中的任意排序组合的集合为元胞空间, 而其中每一个组合 $CellX$ 为元胞。元胞邻居采用扩展 Moore 邻居类型:

$$N_{Moore} = \{CellY | \text{diff}(CellY - CellX) \leq r, CellX, CellY \in L\}$$

其中 $\text{diff}(CellY - CellX) \leq r$ 为两个组合排序的差异, 若无差异为 0, 有差异时, 最小为 2。 r 为差异的程度, 文中 r 为 2。

A 为蚂蚁群体, m 为蚁群算法中全体蚂蚁的数量, $A = \{a_1, a_2, \dots, a_m\}$; d_{ij} 为城市 c_i 和 c_j 之间的距离, 其中 $c_i, c_j \in C$; τ_{ij} 表示在 t 时刻城市 c_i 和 c_j 之间路径连线上残留的信息量。初始时刻各条路径上的信息量相等, 设 $\tau_{ij}(0) = c$ (c 为常数)。蚂蚁群体中蚂蚁 k ($k=1, 2, \dots, m$) 在运动过程中, 根据各条路径上的信息量决定转移的方向, $p_{ij}^k(t)$ 表示在 t 时刻蚂蚁 k 由 i 转移到 j 的概率, 求得一目标解。然后依据元胞邻居的定义计算其邻居的目标解, 比较元胞和其邻居的差异, 选择最好的目标解。

4 基于元胞蚂蚁算法的故障诊断

4.1 诊断方法描述

故障样本的分类与 TSP 问题的需访问的城市相似性。每一组故障样本组合为一元胞, 而所有的样本组合构成元胞空间。

在故障归类过程中, 故障样本之间的近邻函数值越大, 属于不同类的可能性越大。为此, 利用元胞蚂蚁算法, 对全部的故障样本都进行了访问, 形成了全部样本的一个有序的连接, 从中获取近邻函数值最大的两个故障点, 以此分为两个类, 计算其连接总损失; 然后在每个子类中用前面的方法获取近邻函数值最大的两个故障点, 形成新的类, 计算连接总损失; 如此类推, 直到样本数据之间的连接总损失最小为止, 最终形成的故障分类。

获取近邻函数值最大实际上是一种扩展的 TSP 问题, 要求蚂蚁在访问样本点过程中经过的最长距离最长, 其数学模型可写成:

$$\text{Max } Z = \max \{L_{ij} \cdot x_{ij} | i, j \in V\}$$

$$\text{s.t.} \begin{cases} \sum_{j=1}^n x_{ij} = 1, & i \in V \\ \sum_{i=1}^n x_{ij} = 1, & j \in V \\ \sum_{i \in S} \sum_{j \in S} x_{ij} \leq |S| - 1, & \forall S \subset V \\ x_{ij} \in \{0, 1\} \end{cases}$$

其中, L_{ij} 为样本 y_i 和 y_j 之间的近邻函数值。

定义 1 蚂蚁的相邻结点的转移概率定义为:

$$P_{ij}^k = \frac{[\tau_{ij}]^\alpha [L_{ij}]^\beta}{\sum_k [\tau_{ij}]^\alpha [L_{ij}]^\beta}$$

其中, τ_{ij} 为 y_i 和 y_j 间的信息素强度; α —信息素的相对重要性 ($\alpha \geq 0$), β 为近邻函数值的相对重要性 ($\beta \geq 0$)。

定义 2 信息素强度更新方程为:

$$\tau_{ij}^{new} = \rho \cdot \tau_{ij}^{old} + \sum_k \Delta \tau_{ij}^k$$

其中, $\Delta \tau_{ij}^k$ 为蚂蚁 k 于边弧 (i, j) 上留下的单位长度轨迹信息素数量:

$$\Delta \tau_{ij}^k = \begin{cases} Q/Z_k & \text{若 } (i, j) \text{ 在最优路径上} \\ 0 & \text{其他} \end{cases} \quad (Z_k \text{ 为总损失})$$

Q 体现蚂蚁所留轨迹数量的一个常数, ρ —体现强度的持久性 ($0 \leq \rho < 1$)。

定义 3 元胞演化规则

依据元胞邻居的定义计算其邻居的目标解, 比较元胞和其邻居的差异, 选择最好的目标解。

具体算法如下:

(1) 数据准备过程: 获取近邻函数矩阵 L

根据故障样本的分布, 用欧氏距离计算故障征兆样本数据的距离矩阵 Δ 。

利用矩阵 Δ , 构造故障样本的近邻矩阵 M , 其元素 M_{ij} 为样本 y_i 对 y_j 的近邻函数值。

形成近邻函数矩阵 L , 其元素为: $L_{ij} = M_{ij} + M_{ji} - 2$ 。

如果 y_i 和 y_j 间有“连接”关系, 则 L_{ij} 给出了它们之间的近邻函数值。置 L 矩阵的对角元素 L_{ij} ($i=1, 2, \dots, N$) 的值为 $2N$ 或更大。

(2) 使用元胞蚂蚁算法获取最大近邻函数值

每只蚂蚁根据邻域的转移概率决定下一个故障样本数据访问点, 对全部故障样本数据完成了一次访问; 按元胞邻居的定义, 在邻居范围内演化, 记录最好解。

(3) 打断当前有连接情况的近邻函数值最大的样本之间的连接, 从而形成不同的故障分类。计算所有故障类别的类内损失 L_{IA} 和类间损失 L_{IR} 及总损失 J_{NV} 。

(4) 重复步骤 3 获取近邻函数值最大的两个故障点, 计算总损失 J_{NV} , 直到每一只蚂蚁的 J_{NV} 最小, 完成一次故障的识别。

(5) 取所有蚂蚁中 J_{NV} 最小的蚂蚁的识别结果作为本次循环的求解结果。

(6) 调整故障征兆样本数据连接之间的信息素大小。

(7) 判断循环次数是否达到最大步数的给定值, $J_{NV} = L_{IA} + L_{IR}$

是否达到给定的最小值,是则中止运行给出最终解,否则返回步骤2。

4.2 算法仿真

在此以文献[4]实例进行仿真测试,所得结果与文献[4]相同,可将数据分为4类,图1给出了总损失的收敛图形。从图中可见收敛较快,能比较快地达到预定效果。

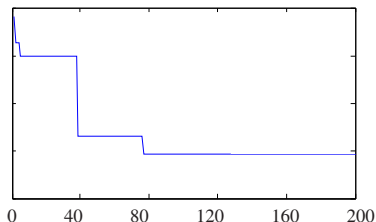


图1 总损失的收敛曲线

5 结束语

故障诊断的准确性和及时性,对系统的安全运行和减少损失有着重要的意义。随着现代工业技术的发展,智能诊断方法不断涌现,得到了飞速发展。提出了一种基于元胞蚂蚁算法的故障归类方法,该算法有利于准确快速地确定故障的原因,为故障诊断提供参考。

参考文献:

[1] 吴今培,肖健华.智能故障诊断与专家系统[M].北京:科学出版社,

(上接 190 页)

- [2] Akira S, Akio S, Hiroyuki A, et al. Dynamic shadow compensation of aerial images based on color and spatial analysis[C]//IEEE International Conference on Pattern Recognition, Barcelona, Spain, 2000:317-320.
- [3] 初秀民,王荣本.基于神经网络的沥青路面破损图像识别研究[J].武汉理工大学学报:交通科学与工程版,2004,28(6):374-378.
- [4] 顾晓东,郭仕德,余道衡.基于 PCNN 的图像阴影处理新方法[J].电子与信息学报,2004,26(3):479-482.

(上接 210 页)

间内以低分辨率数据响应用户,其数据量小,传输时间短,占用带宽小,可避免下载大量的不必要的数据而浪费等待时间和网络带宽,是理想的渐进式传输数据模型。显而易见,该算法对于提高矢量地图数据的网络传输效率具有十分显著的效果。

5 结论

矢量数据的网络渐进式传输,保障了 GIS 信息发布的实时性,缩短了用户等待的时间。提出了一种矢量曲线数据的渐进式传输算法,采用等级搜索树和失高分级算法来实现曲线数据的多分辨率化,是这种传输算法的关键。采用该数据结构便于实现矢量曲线数据的多分辨率增量存储,从而实现矢量地图数据高效的网络渐进式传输,提高了系统的实时性,且极大地节省了带宽。

矢量数据的渐进式传输技术与制图综合技术密切相关,目前还存在着很多问题,如地图矢量数据的动态综合条件下的渐

1997.

- [2] 沈艳霞,纪志成,姜建国.电机故障诊断的人工智能方法综述[J].微特电机,2004,32(2):39-42.
- [3] 姚文俊.基于遗传算法的故障诊断的研究[J].微计算机应用,2004,25(3):280-283.
- [4] 孙京浩,李秋艳,杨欣斌,等.基于蚁群算法的故障识别[J].华东理工大学学报,2004,30(2):194-198.
- [5] 虞和济,陈长征,张省,等.基于神经网络的智能诊断[M].北京:冶金工业出版社,2000.
- [6] 慈建平,张卫民,那荣起,等.基于遗传算法的压缩机故障诊断[J].北京理工大学学报,2001,21(3):310-313.
- [7] 李旭,徐心和.遗传算法在故障诊断专家系统中的应用[J].控制与决策,1998,13(7):378-379.
- [8] 史东锋,屈梁生.遗传算法在故障特征选择中的应用研究[J].振动、测试与诊断,2000,20(9):171-176.
- [9] 左万里.基于多 Agent 的智能故障诊断系统研究[J].计算机与现代化,2003(8):4-6.
- [10] 吴伟蔚,杨叔子.基于智能 Agent 的故障诊断系统研究[J].模式识别与人工智能,2000,13(1):78-82.
- [11] 朱刚,周政新,耿兆丰.远程网络故障诊断系统 MAS 结构模型探索[J].电气自动化,2003,25(4):36-38.
- [12] 朱刚,周政新,耿兆丰.移动代理在发电机绝缘状态智能诊断网络系统中的应用[J].电气自动化,2003,25(5):28-30.
- [13] 朱刚,马良.TSP 的元胞蚂蚁算法求解[J].计算机工程与应用,2007,43(10):79-80.
- [14] 边肇祺,张学工.模式识别[M].北京:清华大学出版社,2000.

- [5] 高建贞,任明武,杨静宇.路面裂缝的自动检测与识别[J].计算机工程,2003,29(12):149-152.
- [6] Wang Bin, Ono A, Muramatsu K, et al. Automated detection and removal of clouds and their shadows from landsat TM images[J].IE-ICE Transaction on Information and Systems,1999,E82-D(2):453-460.
- [7] 王树根,王军利,王爱萍.基于整体变分模型的影像阴影检测算法研究[J].武汉大学学报:信息科学版,2006,31(8):663-667.
- [8] 朱凯军,周焰,兰祖送.基于区域分割的雾天图像增强算法[J].计算机测量与控制,2006,14(5):661-663.

进式传输,相应要素之间的拓扑关系处理及数据安全等方面,都有待于进一步的研究。

参考文献:

- [1] Douglas D H, Peucker T K. Algorithms for the reduction of the number of points required to represent a line or its caricature[J]. The Canadian Cartographer, 1973, 10(2):112-122.
- [2] Bittenfield B P. Transmitting vector geospatial data across the Internet[C]//Berlin: Proceedings of the Second International Conference on Geographic Information Science, 2002, 2478:51-64.
- [3] 吴焕萍,潘熾.GIS 矢量数据的网络渐进式传输研究[J].高技术通讯,2006,16(3):290-295.
- [4] 艾波,艾廷华.矢量曲线数据的“流媒体”传输[J].海洋测绘,2005,25(3):17-20.
- [5] 毋河海.基于多叉树结构的曲线综合算法[J].武汉大学学报:信息科学版,2004,29(6):479-483.